

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-346348  
 (43)Date of publication of application : 14.12.1999

J1002 U.S. PTO  
 09/942717  
 08/31/01

(51)Int.Cl.

H04N 7/04  
 H04N 7/045  
 H04N 5/04

(21)Application number : 10-153022  
 (22)Date of filing : 02.06.1998

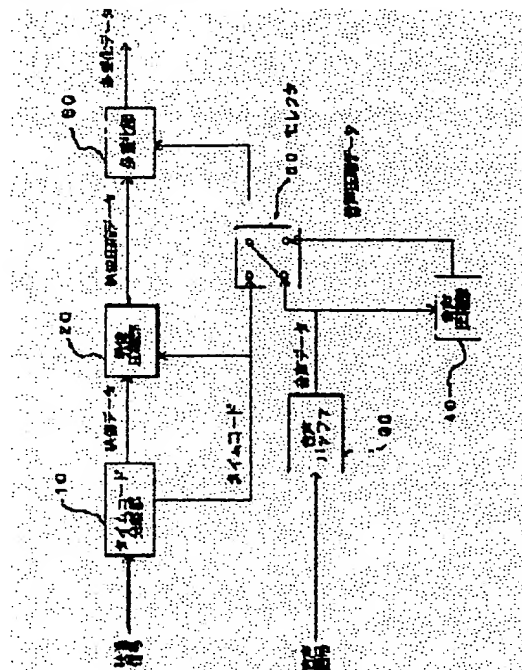
(71)Applicant : NEC CORP  
 (72)Inventor : KATSUBE RYOJI

## (54) VIDEO/AUDIO SYNCHRONIZATION METHOD AND VIDEO/AUDIO MULTIPLEXER

## (57)Abstract:

PROBLEM TO BE SOLVED: To synchronize a video image with audio data in the unit of frames or GOPs.

SOLUTION: This multiplexer is provided with a video compression section 20 that adds a time code included in a video signal to video data to conduct video image compression, a selector 50 that selectively adds a time code included in the video signal to audio data or audio compression data at a position at which synchronization is taken to a frame of the video image or a GOP, and a multiplexer section 60 that multiplexes outputs from the selector 50 and the video compression section 20.



## LEGAL STATUS

[Date of request for examination] 02.06.1998  
 [Date of sending the examiner's decision of rejection]  
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]  
 [Date of final disposal for application]  
 [Patent number]  
 [Date of registration]  
 [Number of appeal against examiner's decision of rejection]  
 [Date of requesting appeal against examiner's decision of rejection]  
 [Date of extinction of right]

Copyright (C); 1998,2000 Japanese Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-346348

(43) 公開日 平成11年(1999)12月14日

(51) Int.Cl.<sup>6</sup>

H 0 4 N 7/04  
7/045  
5/04

識別記号

F I

H 0 4 N 7/04  
5/04

1 0 1  
A

審査請求 有 請求項の数 7 O L (全 6 頁)

(21) 出願番号

特願平10-153022

(22) 出願日

平成10年(1998) 6 月 2 日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目 7 番 1 号

(72) 発明者 勝部 良次

東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

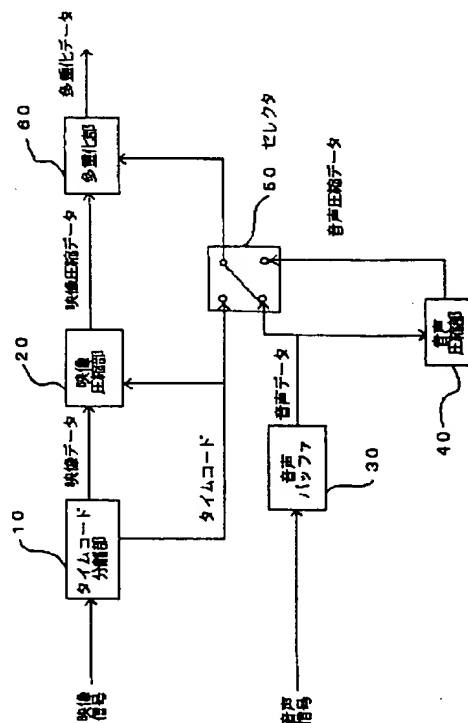
(74) 代理人 弁理士 後藤 洋介 (外 1 名)

(54) 【発明の名称】 映像音声同期方法及び映像音声多重化装置

(57) 【要約】

【課題】 映像のフレーム単位、またはGOP単位で音声データとの同期をとることを可能とする映像音声同期方法及びそれを実現するための装置を提供すること。

【解決手段】 本発明による映像音声多重化装置は、映像信号に含まれるタイムコードを映像データに付加して映像圧縮を行なうための映像圧縮部 20 と、音声データ又は音声圧縮データ上の、映像のフレーム又はGOPと同期をとることの出来る位置に、選択的に映像信号に含まれるタイムコードを付加するセクタ 50 と、セクタ 50 及び映像圧縮部 20 の出力を多重化する多重化部とを備える。



**【特許請求の範囲】**

【請求項 1】 音声データを含む音声信号とタイムコード及び映像データを含む映像信号とを受けて、前記映像データに前記タイムコードを付加して圧縮し、映像圧縮データを生成すると共に、前記音声データに対しても同じタイムコードを前記映像データとの同期をとれる位置に付加してプライベートデータを生成し、当該プライベートデータと前記映像圧縮データとを多重化することにより、再生側において映像データと音声データとの同期をとることの容易な多重化データを生成することを特徴とする映像音声同期方法。

【請求項 2】 音声データを含む音声信号とタイムコード及び映像データを含む映像信号とを受けて、前記映像データに前記タイムコードを付加して圧縮し、映像圧縮データを生成すると共に、前記音声データを圧縮して得られる音声圧縮データに対しても同じタイムコードを前記映像データとの同期をとれる位置に付加し、当該タイムコードの付加された音声圧縮データと前記映像圧縮データとを多重化することにより、再生側において映像データと音声データとの同期をとることの容易な多重化データを生成することを特徴とする映像音声同期方法。

【請求項 3】 前記タイムコードは、映像のフレーム単位に付加されることを特徴とする請求項 1 又は請求項 2 に記載の映像音声同期方法。

【請求項 4】 前記タイムコードは、映像の G O P 単位に付加されることを特徴とする請求項 1 又は請求項 2 に記載の映像音声同期方法。

【請求項 5】 映像データ及びタイムコードを有する映像信号と音声データを含む音声信号とを受けて、映像圧縮を行なう一方、音声データについては非圧縮として、映像音声多重化を行ない多重化データを生成するための映像音声多重化装置であって、  
前記映像信号を受けて、前記タイムコードと前記映像データとを分離すると共に夫々を出力するタイムコード分離部と、  
該タイムコード分離部から前記映像データ及び前記タイムコードを受けて、映像圧縮し、前記タイムコードを含む映像圧縮データを生成するための映像圧縮部と、  
前記音声信号を受けて一時的に格納し、前記音声データを出力するための音声バッファと、  
前記タイムコード分離部及び前記音声バッファに接続され、前記音声データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データの選択を行ない、前記音声データに前記タイムコードの付加されたデータをプライベートデータとして出力するためのセレクトと、  
該セレクトと前記映像圧縮部とに接続され、該セレクトから出力されるプライベートデータと前記映像圧縮データとを多重化して、前記多重化データを生成するための多重化部とを有することを特徴とする映像音声多重化装

置。

【請求項 6】 映像データ及びタイムコードを有する映像信号と音声データを含む音声信号とを受けて、映像圧縮を行なう一方、音声データについても圧縮して、映像音声多重化を行ない多重化データを生成するための映像音声多重化装置であって、

前記映像信号を受けて、前記タイムコードと前記映像データとを分離すると共に夫々を出力するタイムコード分離部と、

該タイムコード分離部から前記映像データ及び前記タイムコードを受けて、映像圧縮し、前記タイムコードを含む映像圧縮データを生成するための映像圧縮部と、  
前記音声信号を受けて一時的に格納し、前記音声データを出力するための音声バッファと、

該音声バッファに接続され、前記音声データを受けて圧縮を行ない、音声圧縮データを生成するための音声圧縮部と、

前記タイムコード分離部及び前記音声圧縮部に接続され、前記音声圧縮データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データの選択を行ない、前記音声圧縮データに前記タイムコードの付加されたデータを出力するためのセレクトと、

該セレクトと前記映像圧縮部とに接続され、該セレクトから出力されるデータと前記映像圧縮データとを多重化して、前記多重化データを生成するための多重化部とを有することを特徴とする映像音声多重化装置。

【請求項 7】 映像データ及びタイムコードを有する映像信号と音声データを含む音声信号とを受けて、映像圧縮を行なう一方、音声データについては圧縮又は非圧縮を選択して、映像音声多重化を行ない多重化データを生成するための映像音声多重化装置であって、

前記映像信号を受けて、前記タイムコードと前記映像データとを分離すると共に夫々を出力するタイムコード分離部と、

該タイムコード分離部から前記映像データ及び前記タイムコードを受けて、映像圧縮し、前記タイムコードを含む映像圧縮データを生成するための映像圧縮部と、  
前記音声信号を受けて一時的に格納し、前記音声データを出力するための音声バッファと、

該音声バッファに接続され、前記音声データを受けて圧縮を行ない、音声圧縮データを生成するための音声圧縮部と、

前記タイムコード分離部、音声バッファ及び前記音声圧縮部に接続され、前記音声データについて非圧縮する場合にあっては、前記音声データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データの選択を行なってプライベートデータを生成し、当該プライベートデータを出力し、前記音声データについて圧縮する場合にあって

は、前記音声圧縮データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データの選択を行ない、前記音声圧縮データに前記タイムコードの付加されたデータを出力するためのセレクトと、

該セレクトと前記映像圧縮部とに接続され、該セレクトから出力されるデータと前記映像圧縮データとを多重化して、前記多重化データを生成するための多重化部とを有することを特徴とする映像音声多重化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、圧縮データに含まれる音声信号と映像信号の同期をとるための方法及びそれを具現化する装置に関し、特に圧縮方式としてMPEG2データを対象とした方法及び装置に関する。

【0002】

【従来の技術】 動画像を構成する映像信号とそれに付随する音声信号とを圧縮するための符号化方式として、MPEG符号化方式が知られている。MPEGとは、Moving Picture (Coding) Experts Groupの略称であり、本来は、カラー動画像を圧縮するための符号化方式の標準化を進める組織の略称であったが、転じて、MPEGにおいて標準化作業が行われている符号化方式の呼称としても使われている。

【0003】 現在、MPEGは、符号化する映像及びそれに伴う音声の形式により、符号化方式をMPEG1、MPEG2、MPEG4の3タイプに分けて標準化が進められている。

【0004】 このうち、MPEG2は、通信や蓄積メディアに限られていた動画像符号化方式であるH.261やMPEG1を、放送、通信、蓄積などのあらゆる分野へ適用できるように発展させた符号化方式であり、転送速度が数Mbpsから数十Mbpsで、テレビ信号に用いられるインタレース(飛び越し走査)信号を含めて、テレビ映像など、高精細テレビ(high definition television, HDTV)を含む高品質の符号化映像を適用対象としている。

【0005】 このMPEG2に従って動画像を構成する映像信号とそれに付随する音声信号とを符号化する際には、映像信号に含まれる画像データ及び音声信号に含まれる音声データは、夫々、次のようにして圧縮される。即ち、画像データは、1画像(フレーム/フィールド)単位でデータ圧縮され、一方、音声データは、24mSEC(音声フレーム)単位でデータ圧縮される。

【0006】 更に、このようにして圧縮された画像データ及び音声データは、夫々が圧縮された後、多重化され、音声を伴う動画像のデータとして、上述したように、放送、通信、及び蓄積メディアなどのあらゆる分野において情報の受け渡しが行なわれ、情報を受け取った

側において、適宜、復号化されることにより、再生される。

【0007】

【発明が解決しようとする課題】しかしながら、MPEG2に従ってデータ圧縮すると、上述したように、映像、即ち画像データに関しては、1画像(フレーム/フィールド)単位で圧縮をかける一方、音声データに関しては、24mSEC単位(音声フレーム)で圧縮をかけており、その上で、圧縮された夫々のデータを多重化するため、圧縮データ上で同期をとろうとしてもデータ数が異なることとなり、結果として、フレーム単位またはGOP単位で映像を編集する場合、音声映像の時間分と一致しないという問題があった。

【0008】 また、タイムスタンプによる同期をとる不連続GOP再生の場合、GOPの先頭で同期をとるため、音声のスキップ、リピートが発生し、忠実な再生ができず、編集処理、特殊再生処理には適していなかった。

【0009】 そこで、本発明は、映像のフレーム単位、またはGOP単位で音声データとの同期をとることを可能とする映像音声同期方法及びその装置を提供することを目的とする。

【0010】 尚、映像データと音声データとの同期を図ることを目的とする従来技術としては、特開平4-285769号公報(以下、従来例1という)、特開平2-135993号公報(以下、従来例2という)、及び特開平8-172597号公報(以下、従来例3という)に開示されているものが挙げられるが、いずれもMPEG2符号化方式特有の問題点について何等言及するものではない。特に、従来例1に関しては、技術内容が古く、後述する本発明とはかけ離れたものであり、また、従来例2及び従来例3については、その構成上、後述する本発明を理解されれば明らかになるように、本発明と比較して、冗長なハードウェアを必要とするものである。

【0011】

【課題を解決するための手段】上述した課題を解決するにあたって、本発明の発明者は、MPEG2符号化方式における特徴の一つである、映像データに対してタイムコードを付加できる点に着目した。その上で、音声データを非圧縮の状態に圧縮された映像データに多重化して伝送する音声非圧縮データ伝送の場合と、音声データを圧縮した状態で圧縮された映像データに対して多重化して伝送する音声圧縮伝送の場合とに分けて、映像データに対して付加したタイムコードを有効に利用することにより、上述した課題を解決することができることを見出した。即ち、本発明による第1の映像音声同期方法として、音声データを含む音声信号とタイムコード及び映像データを含む映像信号とを受けて、前記映像データに前記タイムコードを付加して圧縮し、映像圧縮データ

を生成すると共に、前記音声データに対しても同じタイムコードを前記映像データとの同期をとれる位置に付加してプライベートデータを生成し、当該プライベートデータと前記映像圧縮データとを多重化することにより、再生側において映像データと音声データとの同期をとることの容易な多重化データを生成することを特徴とする映像音声同期方法が得られる。

【0012】また、本発明によれば、第2の映像音声同期方法として、音声データを含む音声信号とタイムコード及び映像データを含む映像信号とを受けて、前記映像データに前記タイムコードを付加して圧縮し、映像圧縮データを生成すると共に、前記音声データを圧縮して得られる音声圧縮データに対しても同じタイムコードを前記映像データとの同期をとれる位置に付加し、当該タイムコードの付加された音声圧縮データと前記映像圧縮データとを多重化することにより、再生側において映像データと音声データとの同期をとることの容易な多重化データを生成することを特徴とする映像音声同期方法が得られる。

【0013】尚、上記映像音声同期方法において、前記タイムコードは、映像のフレーム単位に付加されることとしても良いし、映像のGOP単位に付加されることとしても良い。

【0014】更に、これらの方法を実現するための具体的な手段は、以下に示される。

【0015】即ち、本発明によれば、第1の音声映像多重化装置として、映像データ及びタイムコードを有する映像信号と音声データを含む音声信号とを受けて、映像圧縮を行なう一方、音声データについては非圧縮として、映像音声多重化を行ない多重化データを生成するための映像音声多重化装置であって、前記映像信号を受けて、前記タイムコードと前記映像データとを分離すると共に夫々を出力するタイムコード分離部と、該タイムコード分離部から前記映像データ及び前記タイムコードを受けて、映像圧縮し、前記タイムコードを含む映像圧縮データを生成するための映像圧縮部と、前記音声信号を受けて一時的に格納し、前記音声データを出力するための音声バッファと、前記タイムコード分離部及び前記音声バッファに接続され、前記音声データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データを選択を行なってプライベートデータを生成し、当該プライベートデータを出力するためのセレクトと、該セレクトと前記映像圧縮部とに接続され、該プライベートデータと前記映像圧縮データとを多重化して、前記多重化データを生成するための多重化部とを有することを特徴とする映像音声多重化装置が得られる。

【0016】また、本発明によれば、第2の映像音声多重化装置として、映像データ及びタイムコードを有する映像信号と音声データを含む音声信号とを受けて、映像

圧縮を行なう一方、音声データについても圧縮して、映像音声多重化を行ない多重化データを生成するための映像音声多重化装置であって、前記映像信号を受けて、前記タイムコードと前記映像データとを分離すると共に夫々を出力するタイムコード分離部と、該タイムコード分離部から前記映像データ及び前記タイムコードを受けて、映像圧縮し、前記タイムコードを含む映像圧縮データを生成するための映像圧縮部と、前記音声信号を受けて一時的に格納し、前記音声データを出力するための音声バッファと、該音声バッファに接続され、前記音声データを受けて圧縮を行ない、音声圧縮データを生成するための音声圧縮部と、前記タイムコード分離部及び前記音声圧縮部に接続され、前記音声圧縮データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データを選択を行ない、前記音声圧縮データに前記タイムコードの付加されたデータを出力するためのセレクトと、該セレクトと前記映像圧縮部とに接続され、該セレクトから出力されるデータと前記映像圧縮データとを多重化して、前記多重化データを生成するための多重化部とを有することを特徴とする映像音声多重化装置が得られる。

【0017】更に、本発明によれば、第3の映像音声多重化装置として、映像データ及びタイムコードを有する映像信号と音声データを含む音声信号とを受けて、映像圧縮を行なう一方、音声データについては圧縮又は非圧縮を選択して、映像音声多重化を行ない多重化データを生成するための映像音声多重化装置であって、前記映像信号を受けて、前記タイムコードと前記映像データとを分離すると共に夫々を出力するタイムコード分離部と、該タイムコード分離部から前記映像データ及び前記タイムコードを受けて、映像圧縮し、前記タイムコードを含む映像圧縮データを生成するための映像圧縮部と、前記音声信号を受けて一時的に格納し、前記音声データを出力するための音声バッファと、該音声バッファに接続され、前記音声データを受けて圧縮を行ない、音声圧縮データを生成するための音声圧縮部と、前記タイムコード分離部、音声バッファ及び前記音声圧縮部に接続され、前記音声データについて非圧縮する場合にあっては、前記音声データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データを選択を行なってプライベートデータを生成し、当該プライベートデータを出力し、前記音声データについて圧縮する場合にあっては、前記音声圧縮データに対して、前記映像データと前記音声データとの同期がとれる位置に、前記タイムコードを挿入するように、データを選択を行ない、前記音声圧縮データに前記タイムコードの付加されたデータを出力するためのセレクトと、該セレクトと前記映像圧縮部とに接続され、該セレクトから出力されるデータと前記映像圧縮データとを多重化して、前記多重化データを生成するための多

重化部とを有することを特徴とする映像音声多重化装置が得られる。

【0018】

【発明の実施の形態】以下に、本発明の実施の形態について図面を用いて説明する。本実施の形態による映像音声同期方法は、図1に示されるような映像音声多重化装置上において実現される。実施の形態による映像音声多重化装置は、図1に示されるように、タイムコード分離部10、映像圧縮部20、音声バッファ30、音声圧縮部40、セクタ50、及び多重化部60とを備えている。

【0019】タイムコード分離部10は、映像信号を受けて、当該映像信号に含まれるタイムコードを検出し、映像信号に含まれる映像データとタイムコードとに分離して出力する。

【0020】映像圧縮部20は、タイムコード分離部10から映像データとタイムコードとを受けて、当該映像データに対してMEPG2に従ったデータ圧縮を行ない、映像圧縮データを生成し、出力する。

【0021】音声バッファ30は、音声信号を受けて、一時的に格納し、音声データとして出力するためのものであり、音声圧縮部40は、音声バッファ30から音声データを受けて、MEPG2に従ったデータ圧縮を行ない、音声圧縮データを生成し、出力するものである。

【0022】セクタ50は、タイムコード分離部10、音声バッファ30、及び音声圧縮部40に接続され、タイムコード、音声データ、音声圧縮データのうちのいずれか一つを選択して、出力するためのものである。

【0023】多重化部60は、映像圧縮部20とセクタ50とに接続され、映像圧縮部20からの映像圧縮データとセクタ50により選択されたデータとを多重化して多重化データとして出力するためのものである。

【0024】このような構成を備える本実施の形態による映像音声多重化装置は、次のようにして、映像音声多重化処理を行なう。

【0025】まず、音声データを圧縮せずに映像圧縮データと多重化して多重化データを出力する場合について説明する。

【0026】映像信号は、タイムコード分離部10に入力されると、当該映像信号に含まれるタイムコードを検出され、当該タイムコードと映像データとの分離が行なわれる。タイムコード分離部10は、検出分離したタイムコードと映像データとを夫々別個独立して後段へ出力する。

【0027】映像圧縮部20は、タイムコード分離部10により検出分離された映像データ及びタイムコードの双方を受けて、映像圧縮し、タイムコードを含む映像圧縮データを生成する。

【0028】一方、音声バッファ30は、音声信号を受

けて一時的に格納し、映像のフレーム又はGOP毎に音声データとして出力する。

【0029】セクタ50は、音声バッファ30から映像のフレーム又はGOP毎に出力される音声データのヘッダに対して、タイムコード分離部10からのタイムコードを付加挿入するようにして、データの切替えを行い、音声データに対して映像のフレーム又はGOP毎にタイムコードの付加されたデータをプライベートデータとして、出力する。この際、音声圧縮部40は、稼動していてもしていなくても良い。

【0030】セクタ50から出力されたプライベートデータは、映像圧縮部20から出力される映像圧縮データと、多重化部にて多重化され、多重化データとして出力される。

【0031】このように本実施の形態によれば、音声データを圧縮せずに映像圧縮データと多重化して多重化データを出力する場合には、音声データが、映像フレーム又はGOP毎にバッファリングされると共に、映像フレームと同じタイムコードを付けられて、これをプライベートデータとして映像圧縮データと一緒に多重化され、多重化データが生成されることから、この多重化データを再生したり、編集したりする側においては、映像のフレーム単位又はGOP単位で音声データとの同期をとった場合であっても、音声のスキップ及びリピートの発生が生じないと共に、音声と映像の時間分が一致しないといった問題の防止が図られることとなる。

【0032】次いで、音声データを圧縮すると共に映像圧縮データと多重化して多重化データを出力する場合について説明する。

【0033】映像信号は、タイムコード分離部10に入力されると、当該映像信号に含まれるタイムコードを検出され、当該タイムコードと映像データとの分離が行なわれる。タイムコード分離部10は、検出分離したタイムコードと映像データとを夫々別個独立して後段へ出力する。

【0034】映像圧縮部20は、タイムコード分離部10により検出分離された映像データ及びタイムコードの双方を受けて、映像圧縮し、タイムコードを含む映像圧縮データを生成する。

【0035】一方、音声バッファ30は、音声信号を受けて一時的に格納し、映像のフレーム又はGOP毎に音声データとして出力する。尚、この音声データは、音声圧縮部40及びセクタ50の双方に出力されるが、本例においては、セクタ50での選択はなされない。

【0036】音声バッファ30から音声データを映像のフレーム又はGOP単位に受けた音声圧縮部40は、音声圧縮を行ない、音声圧縮データとしてセクタ50に対して出力する。

【0037】セクタ50は、音声圧縮部40からの映像のフレーム又はGOP毎に圧縮された音声圧縮データ

を受けて、その音声圧縮データのヘッダに対して、タイムコード分離部 10 からのタイムコードを付加挿入するようにして、データの切替えを行い、タイムコードの付加された音声圧縮データを出力する。

【0038】セクタ 50 から出力されたタイムコード付きの音声圧縮データは、映像圧縮部 20 から出力される映像圧縮データと、多重化部にて多重化され、多重化データとして出力される。

【0039】このように本実施の形態によれば、音声データを圧縮すると共に映像圧縮データと多重化して多重化データを出力する場合には、音声データが、映像フレーム又は GOP 毎にバッファリングされて音声圧縮されると共に、その音声圧縮データのヘッダに対して映像フレームと同じタイムコードを付けられて、これを映像圧縮データと一緒に多重化され、多重化データが生成されることから、この多重化データを再生したり、編集したりする側においては、映像のフレーム単位又は GOP 単位で音声データとの同期をとった場合であっても、音声のスキップ及びリピートの発生が生じないと共に、音声と映像の時間分が一致しないといった問題の防止が図られることとなる。

【0040】尚、本実施の形態においては、上述した説明から理解されるように、セクタ 50 により音声バッファ 30 からの出力と音声圧縮部 40 からの出力とを場合に応じて選択することにより、音声データを圧縮する

と共に映像圧縮データと多重化して多重化データを出力する場合及び音声データを圧縮せずに映像圧縮データと多重化して多重化データを出力する場合の双方に対応することとしたが、夫々の場合用の装置として、別個に構成することとしても良いことは、言うまでもない。

#### 【0041】

【発明の効果】本発明によれば、映像信号に含まれるタイムコードを利用して、映像データに対して当該タイムコードを付加して圧縮すると共に、同じタイムコードを音声データ又は音声圧縮データ上の映像のフレーム又は GOP に対応する位置にヘッダとして挿入付加することにより、映像のフレーム単位、または、GOP 単位に音声のデータを完全に同期をとることが可能となり、映像音声とも編集可能となる。

#### 【図面の簡単な説明】

【図 1】本発明の一実施の形態による映像音声多重化装置の構成を示すブロック図である。

#### 【符号の説明】

10	タイムコード分離部
20	映像圧縮部
30	音声バッファ
40	音声圧縮部
50	セクタ
60	多重化部

【図 1】

